Remarks

Claim Rejections under 35 USC 112

Claims 1 and 10 were rejected under 35 USC 112, first paragraph as failing to comply with the written description requirement. When claims 1 and 10 were amended, a new term "integrator" was introduced to describe a previous component ("router" col. 1, paragraph [0008]) for the purpose of further clarification and better defining the Applicant's invention. The Applicant requests withdrawal of the terms "integrator", "integrating" and "integrated" from Claims 1 and 10 to comply with written description requirements and to allow claims 1 and 10 based on these amendments.

Claim Rejections under 35 USC 103

Claims 1 and 10 were rejected under 35 USC 103(a) as being unpatentable over Rondel et al. (USPN 4,984,177) referred to as Rondel hereinafter in view of Meredith (US Patent No. 5,796,916) referred to as Meredith hereinafter. The Applicant requests allowability of Claims 1 and 10 based on the following:

There are fundamental and distinct differences between the device and method described in Rondel et al. (USPN 4,984,177) for language translation and Rousseau's device and method (US2004/0122678 A1), specifically, the Rondel device is referenced as a training tool that functions as a translator 'only' after it has been trained to the voice of a user (col. 1, lines 59-61). Further to this point, the level of training that is required for the Rondel device to function properly depends on the degree of speaker independence in the voice recognition circuit (col. 1, 61 - 63). An objective of the Rousseau design is for multiple users and speakers to have the ability to engage in real-time language translation to effect conversation between the users.

Additionally, the device described by Rondel determines the nature (content component), of the spoken words or sentence (col.2, lines 6-9)i.e., what is being said by the speaker by comparing the output of the voice recognition circuit with the voice pattern, and using this comparison of the spoken words or phrases with the voice pattern in order to locate the equivalent translations (col 2, lines 9 - 11). This requirement necessitates the content of what is spoken be limited to a

specific format, that is, 'sentences in the form of instructions', questions that are translatable' and 'preferably each question or instruction end with a specific word unrelated to the content such as PLEASE to comprise strings of words and phrases that make up sentences that are combinable only in a logical manner' and illogical combinations of words or phrases are rejected as allowable combinations of words and phrases for which voice patterns have been stored (col.2, lines 7- 18). The device described by Rondel not only requires logical string combinations in the form of instructions or questions requiring YES, NO as specific format during the input of a speaker's voice , but preferably each instruction or question ends with a specific word that is tonally and contextually unrelated to the instruction or question such as PLEASE (col 2. lines 31-37). An ending term is a requirement of the Rondel translation process to let the translator know when the string combination to be translated has ended and translation is to begin (col. 2, lines 37-40).

By comparison, the Rousseau invention describes a device that functions as a translator and uses voice recognition technology that doesn't require users to pre-train the device to the speaker's voice patterns in order to function properly. Nor does the Rousseau invention require the user's

speech to be pre-determined and formatted as a string of words and phrases that are combinable only in a logical manner and ending each sentence with unrelated words to signal the end of the phrase to be translated and trigger the beginning of the translation process. The primary function of the Rousseau invention is as a translation device that has the ability to operate in real- time with multiple users, not as a training tool that is used (by Rondel) to perfect speech patterns in a second language; and as such the Rousseau method is designed to utilize voice analysis technologies within the device to analyze the inputs of the speakers in detail in order to create voice prints that are unique to each individual as well as have the capability to process this voice print data with the translation to affect the quality of the translated output to approximate each of the speaker's voice characteristics. This feature is unique to Rousseau and lends itself not only for use as a translation device but as a conversation tool for multiple users.

Meredith (USPN 5,796,916) discloses a TTS (text-to-speech) method and apparatus for prosidy for synthetic speech prosody determination (Figure 2) where user sound input is analyzed for pitch and aligns existing phonetic (text) transcriptions with pitch contour to output a more natural sounding synthesized voice. Meredith further discloses that the objective of this text-to-speech method is to edit, correct (col. 2, lines 54 - 64) and improve the intonation (only) of synthesized speech through alignment against a recorded natural utterance and as a method for intonation specification by aligning voicing sections of a natural utterance to one or more voicing sections of a phonetic text stream and applying intonation of one or more voicing sections of the natural utterance against the voicing sections of the phonetic stream (col. 3, lines 4 - 10).

Date: 9/18/2008

The Rousseau invention employs a very different technology than Meredith, being a STS (speech -to-speech) device and synthesis method, rather than a TTS (text-to-speech) technology. The Rousseau' embodiment's objective is to make use of a software model that generates a more of an in-depth analysis of the components of the voice (voice printing) in a speech- to- speech system to assign quantitative values to the voice patterns and generate detailed inputs of the speaker's voice, allowing this data to be converted into 3-D predictive analytic simulations that fingerprint and map the user's voice patterns for components such as pitch (intonation), frequency, volume, etc., much like body movement is characterized and simulated in 3-D to predict fit and wear of clothing and products or to analyze vibration in inanimate objects. The voice specifications of the Rousseau model are transmitted through the routers and are combined with the translated output using predictive responses that allow the integration of real-time data (Rousseau's invention) versus voice section alignments and lowest accumulated error techniques employed in text-to-speech methods used in Meredith with the objective of both technologies (Rousseau and Meredith) to provide a more natural sounding output but with Rousseau's invention expanding further, to replicate the unique characteristics of the speaker's (human) voice i.e., provide translations with the speaker's voice characteristics and translations that are geared to match the technical specifications of a speaker's voice with enough accuracy that recipients of the translations can recognize which user's voice is being projected in the translation

even though the language of the content component has changed. This objective of the Rousseau

device is an original objective and is believed to be of importance as an evolution in the next

generation translation devices and technologies that seek to improve the continuity and quality of translational exchange between parties through advanced voice technologies, software and machines that are able to bridge many of the gaps noted with translations devices today; to effect real-time translations and conversation between individuals and while maintaining the human components in translation and language exchange processes to personalize the language translation between parties. The technique employed in Meredith on the other hand, is applied to a synthetic (robotic) sound model to edit intonation only and provide a more natural sounding synthesized model.

Claims 1 and 10:

Rondel and Meredith are distinctively different technologies from the device and method described by Rousseau and should not be used as the basis for rejection of the Rousseau invention. Rondel, Meredith and Rousseau are all distinctive technologies, and the following descriptions are necessary features and anticipated functions for translation devices and methods used by those skilled in the art and would not lead to rejection as unpatentable solely on the basis of expected similarities.

i. receiving a speech input signal in a first language, ii., converting the speech input signal into a digital format comprising a voice model component representing a speech pattern of the speech

input signal and content component representing a content of the speech input signal, iii, translating the content component and iv., generating an audible output signal comprising the translated content . All of these are functions that are necessary features for translation devices and would be anticipated in the disclosures of these devices, methods and languages used with different technologies for translation and as such should not lead to rejection of Rousseau's method as unpatentable.

(Fig 2) of the Meredith disclosure discusses alignment of text input (phonetic transcription) with pitch contour and a text-to-speech method that specifically improves intonation of a pre-existing synthetic speech model to provide a more natural sounding speech synthesis. Although similar in the objective to provide a more natural sounding synthesis during translation, text-to-speech is a significant distinction with Meredith and is not the same technology as the speech-to-speech method disclosed by Rousseau for a similar purpose.

Claim 2 : Identifying the first language from the content component is a feature that is sometimes used in translation systems to specify the target language. The Rousseau method uses voice recognition technology to identify the first language (speakers language) from the content component and is not designed as a single user translation device (like Rondel).

Claim 3: Predetermination of the target (first speakers) language is an obvious feature that is required in the use of translation devices and methods, however, the Rousseau invention does not rely on the use of cartridges for pre-determining the first language (as Rondel); instead the first language is retrieved in the translation process and generation of the translated outputs is either a function of the voice recognition technology (claim 2) or predetermined language as claim 3 indicates. Technologies are available today with the capability to support either choice, and this focus is in the scope of the Rousseau method for shifting the language exchanges from mere translations to exchanges that are more conversational in nature and can be used with multiple speakers and parties.

Claim 4: The voice model component that comprises a point cloud as disclosed in Rousseau, claim 4, is intended as a set of collective data points that are useful as a visual, more comprehensive representation to map the details and characteristics of the speaker or a multitude of speaker's voice patterns and is not the same as a digitally encoded voice pattern as mentioned in Rondel (col. 9, lines 40-45). A point cloud can provide a comprehensive 3-dimensional view and the ability to understand the areas of data overlap and relationship of multiple voice patterns such as where voice characteristics of multiple users would converge and provide similar sound traits. The voice model component that is represented 3-dimensionally with the use of a point cloud data set offers greater detail and is broader in scope than 1 or 2 dimensional graphical representations. A point cloud that is generated by FEA or finite difference techniques can be digitized and streamlined as smaller units of data and information to save computer space and improve the transmittance of data to storage and other banks of data in the translation system's computer .

Claim 5: The voice model component that is represented by a point cloud in Rousseau, claim 5, is a further abbreviated set(s) of collective data points that are useful as a visual representation and helpful in mapping the details and characteristics of the speaker's voice patterns.

This set of data is not created or presented in the same form as a digitally encoded voice pattern that is mentioned in Rondel (col 9, lines 39 - 43). The voice model component that is mapped as an abbreviated point cloud or data set can be digitized and streamlined as smaller units of data

and information as a digitally encoded pattern, if needed. A point cloud with an abbreviated

dimensional scope versus that of a standard analysis allows the user an abbreviated snap shot representation and is of value when the user needs to expedite the manipulation of data or further improve the speed of data downloads, converting and transmittance in a system's computer.

Computer tools such as Solid Works programs can be used to build abbreviated and full version point clouds, but the process tends to be technique dependent, more labor intensive and the data sets don't offer the same degree of technical integrity.

Claim 6: The content component as described in Rousseau, claim 6, would typically comprise a sentence, a cluster or more than one word, but would involve at least one word for the voice recognition and translation process. The one word as described in this claim is not intended as a means of triggering a translation system to realize the target language as is indicated by the use of the term "utterance, French" in Rondel (col. 9, lines 28 – 30) to realize what type of language is intended for the target language. Voice recognition software and technologies available in today's computers are capable of triggering the language of a speaker with minimal input (content component) from the speaker such that the use of cartridges and pre-selection of the first language (predetermination) prior to translation are often not needed. Translation devices and processes that rely on software technologies to determine the first language (like Rousseau) typically do not necessitate the use of cartridges to pre-determine the first language of the speaker as indicated by the Rondel method in this claim.

Claim 7: As an older technology, Rondel (1991) describes both a process for determining what was spoken by the user and which words or phrases are to be translated by comparing digitally coded voice patterns. More recent technologies (Rousseau) make use of digitized signals to improve the speed of the translation and use voice patterns to significantly improve the quality of the translated output (versus) voice training, to improve conversation between parties.

Claim 8: The objective of Rousseau's device and method is the capability to produce quality translations and technical (sound) matches to the speaker's voice and achieve a translated voice pattern that sounds like the speaker's voice by characterizing the speaker's voice and the production of a quality voice print incorporating the speaker's voice patterns and combining the elements of this data set with the translated (content component) data set to produce like-sounding voice patterns in the translation.

The quality of the translation and how well it replicates the speaker's voice characteristics is contingent on the quality of the voice print. And the ability to drive translations (words) that are not affected by the dialect of the speaker. Since this principle is based on predictive math,
translations in the second language that sound like the speaker (to include a speaker's dialect) can also be predicted using these simulations.

Claim 9: Real-time communication is a feature that is desired for translation devices to improve conversation between parties and provide a favorable interactive experience.

Rondel and Rousseau both include real-time communication specifications, however, the technologies and objectives of Rondel and Rousseau remain significantly different and this premise is further demonstrated by the fact that real-time communication and operation of the Rondel's device is only possible when the device has been sufficiently trained to the user's voice.

Date: 9/18/2008

Rousseau's technology is operable without training and does not require training to the speaker's voice to affect proper function of the device.

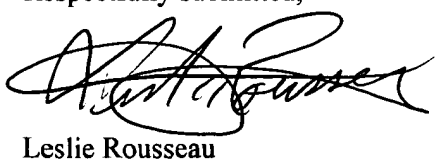
The Applicant respectfully requests examination and allowance of the invention discussed herein, Rousseau's Device and Method for Translating Language (US2004/1022678 A1),

based on the merits of the invention, the amended claims and explanations provided pertaining

to Claim Rejections under 35 USC 112 and 35 USC 103.

Please feel free to contact Leslie Rousseau, the undersigned Applicant at 303 -750- 4205 or widdie8@msn.com, if the Examiner has questions or would like to speak with me directly concerning this filing.

Respectfully submitted,



Leslie Rousseau
1100 South Moline Street
Aurora, CO 80012
303-750-4205
widdie8@msn.com